MONTHLY WEATHER REVIEW

CHARLES F. BROOKS, Editor.

VOL. 47, No. 12.
W. B. No. 702.

DECEMBER, 1919.

CLOSED FEB. 4, 1920
ISSUED MAR. 10, 1920

A STATISTICAL STUDY OF WEATHER FACTORS AFFECTING THE YIELD OF WINTER WHEAT IN OHIO.

By THOMAS ARTHUR BLAIR, Meteorologist.

[Dated Weather Bureau, Salt Lake City, Utah, Dec. 1, 1919.]

SYNOPSIS.

The statistical method is applied to the problem of determining what are the important weather factors affecting the growth of winter wheat in Ohio, and their relative importance. The results are expressed as partial correlation coefficients and in linear regression equations of the form,

$$Y=a+b_1x_1+b_2x_2+b_3x_3+b_4x_4+ \quad \ldots.,$$

in which the coefficients are evaluated by least square methods.

Because of the difficulty of securing extensive data for other weather elements, it is necessary to deal chiefly with temperature and precipitation values. In general, it is to be expected, because of the relatively large and well-distributed rainfall of Ohio, that temperature variations will have more influence than precipitation variations upon the yield.

For the State as a whole, correlations of monthly weather values with the "condition" reports of the Bureau of Crop Estimates, and with the reported yields, show no very close relationships. The correlations with condition give a general indication that a wet autumn, a warm and dry winter and spring, especially a warm March, and a cool and wet May are the most favorable weather conditions. Yield correlations suggest a warm March and June and a cool and dry May as the only important requisites for a good yield.

In Fulton County in northwestern Ohio, and in three counties in the central part of the State, certain 10-day periods in April, May, and June are found to exert a more effective influence on the yield than all other weather conditions combined, except that in Fulton County the March snowfall is also an important factor. It is weather conditions during these 10-day periods, especially temperature conditions, that largely determine yield. These periods are connected especially with the jointing, heading, and filling stages in the growth of the plant.

INTRODUCTION.

In addition to being one of the oldest of cultivated crops, wheat is probably the most important, as world events of the past few years have sharply emphasized. While the climatic zones in which it can be grown successfully are well recognized and the cultural practices in handling it are pretty firmly established, and though there has grown up a considerable body of traditional or empirical knowledge or assumption concerning the influence of the weather factors, yet the actual effects of various kinds of weather upon the progress and yield of the crop are only very imperfectly known. The following study is an attempt to determine more definitely what are the major weather controls in the growth of winter wheat in Ohio, and their relative importance.

There are two general methods by which such a problem may be attacked. One is the experimental method of planting the grain in plots under more or less controlled conditions. This has certain great advantages, but can practically be carried on only at an agricultural experiment station under a settled and continuous policy for several years. The other plan of attack, and the one employed herein, is the statistical method, in which the actual yields under commercial conditions are compared in historical series with the recorded weather conditions. Where reliable records are available for a considerable period this method seems to offer a valuable field of work, supplementary to and in some respects superior to the experimental method, though it must necessarily omit many details, such as differences of culture, vitality of seed, time of planting, kind and condition of soil, etc., and deal only with averages and large factors. It must overlook entirely minor factors, as well as those which, though they may be of first importance in a particular plot, are not generally applicable over large areas, and for this reason offers, perhaps, a better opportunity of viewing the larger and more general controls.

METHODS OF COMPUTATION.

Two particular methods of computing the relationships between weather and yield have been used in this paper. One is the method of partial correlation coefficients, as developed in the textbooks on the theory of statistics,[1] and as previously exemplified in the MONTHLY WEATHER REVIEW.[2] The other plan has been to develop and evaluate a linear regression equation by means of which the yield is expressed as a function of from three to six weather elements. The general equation is written thus:

$$Y=a+b_1x_1+b_2x_2+b_3x_3+b_4x_4+\ \ldots. \qquad (1)$$

In which Y is the yield; x_1, x_2, x_3, x_4, ... are the various weather elements, such as mean temperature, total precipitation, percentage of sunshine, expressed numerically; a, b_1, b_2, b_3, b_4, ... are numerical quantities having a constant value for a given equation, to be determined from the data. The assumptions of this equation are similar to those involved in the correlation calculation, first, that a straight line relation subsists between the yield and the weather (which is doubtful in cases of extreme weather conditions), and, second, that the most important influences have been embodied in the particular equation, or that the influences not included have varied independently of those included.

The solution is by means of normal equations as developed in the theory of least squares. The equation may first be simplified by the elimination of a, as follows: The normal equation for a, obtained by multiplying each of the observation equations by the coefficient of a, in this case unity, and taking their sum, will be of the form

$$\Sigma Y=na+\Sigma b_1x_1+\Sigma b_2x_2+\Sigma b_3x_3+\Sigma b_4x_4+\ \ldots. \qquad (2)$$

from which

$$a=\frac{\Sigma Y}{n}-\frac{\Sigma b_1x_1}{n}-\frac{\Sigma b_2x_2}{n}-\frac{\Sigma b_3x_3}{n}-\frac{\Sigma b_4x_4}{n}-\ \ldots. \qquad (3)$$

[1] Yule, Introduction to the Theory of Statistics. [The Computer's Handbook, Section V, Meteorological Office, Great Britain.]
[2] Blair, T. A. Partial Correlation Applied to Dakota Data on Weather and Wheat Yield. MONTHLY WEATHER REVIEW, Feb. 1918, 46: 71-73.

162722—20——1

Substituting this value in the original equation gives—

$$Y-\frac{\Sigma Y}{n}=\left(x_1-\frac{\Sigma x_1}{n}\right)b_1+\left(x_2-\frac{\Sigma x_2}{n}\right)b_2+\left(x_3-\frac{\Sigma x_3}{n}\right)b_3+\left(x_4-\frac{\Sigma x_4}{n}\right)b_4+\ \ldots\ . \quad (4)$$

Now it will be noticed that the expressions, $Y-\frac{\Sigma Y}{n}$, $x-\frac{\Sigma x_1}{n}$, etc., are in each case the differences between the values for the individual years and the average values of the same quantities, that is to say, are the departures from the means. Designating these departures by y, D_1, D_2, D_3, D_4, ..., we have for the final form of the equation [3]

$$y=b_1D_1+b_2D_2+b_3D_3+b_4D_4+\ \ldots\ . \quad (5)$$

From equation 5 the normal equations become—

$$\Sigma D_1y=b_1\Sigma D^2_1+b_2\Sigma D_1D_2+b_3\Sigma D_1D_3+\ \ldots\ .$$
$$\Sigma D_2y=b_1\Sigma D_1D_2+b_2\Sigma D^2_2+b_3\Sigma D_2D_3+\ \ldots\ .$$
$$\Sigma D_3y=b_1\Sigma D_1D_3+b_2\Sigma D_2D_3+b_3\Sigma D^2_3+\ \ldots\ .$$

It is then necessary to prepare tables of the data used, together with the departures, squares of departures, and products of departures appearing in the normal equations, and then to solve these equations simultaneously for the values of b_1, b_2, b_3, ... The solution is most conveniently performed, especially where the number of equations exceeds three, by writing the required values in the form of determinants, and reducing these by successive steps to the second order before expanding. The mechanical work of solution increases very rapidly of course when the unknowns increase beyond four or five.

Some of the advantages of this method over that of partial correlation coefficients are: The effect of the several factors considered is expressed directly in an equation from which the yield for any year can be calculated. The numerical work can be quickly and absolutely checked by substituting in the normal equations. The accuracy of the solution as a forecasting equation can be expressed mathematically by calculating the "scatter"[4], i. e., the square root of the mean square deviation of the calculated from the recorded yields. On the other hand, an inspection of the equation, without this added labor of calculating the results for individual years, will give no hint of its accuracy, whereas the partial correlation coefficients show at once by the nearness of their approach to unity whether the important factors have been included and their relative importance. In the linear equation the numerical values are the result of the units used, which are variously inches, degrees, percentages, etc., and to obtain the relative importance of the coefficients, b_1, b_2, b_3, ..., the unit must be eliminated by multiplying each by its standard deviation, σ. The values of $b_1\sigma_1$, $b_2\sigma_2$, $b_3\sigma_3$, ... are given in the following discussion, accordingly, for the purpose of showing the relative values of the various factors.

In general, the method of partial correlation has been found most satisfactory and convenient for discovering the important factors and their relative weights, while the regression equation gives a more elegant and complete expression of the final relationships and of their actual value in accounting for the yield. The shortness of the record for statistical purposes combined with the complexity of the factors and the long period during which the plant is exposed to the weather renders it impracticable to make preliminary determinations by such means as dot charts and total correlations of single factors. Such results are unreliable, and in order to reach trustworthy conclusions it is necessary to combine a number of elements in one calculation, and to include those elements showing large correlations, whether obviously related, such as temperature and snowfall of March, or apparently unrelated, as in the case of temperatures or precipitation amounts for different periods.

GENERAL CONSIDERATIONS.

Winter wheat in Ohio is planted in September and harvested in July, and is therefore subject to the influence of the weather for nine months. It would not be expected *a priori*, therefore, that any one month or short period, or any one series of weather events, unless extremely unusual, would have a predominating influence, but rather that favorable or unfavorable conditions in each of the months between planting and harvest would show in the final yield. This is the common belief, countenanced by the practice of the Bureau of Crop Estimates of issuing estimates of condition as early as December 1. The conclusions of this paper cast doubt upon the truth of this belief and the value of this practice as applied to the early stages of growth.

Ohio has a mean annual precipitation of about 38 inches, very well distributed through the year, each month having a normal precipitation of more than 2.50 inches, but winter wheat is successfully grown in interior California, having a normal annual precipitation of 12 inches, and in central and western Kansas with from 26 to 19 inches. The least annual amount ever recorded in Ohio is 28.46 inches, and the months with less than 1 inch are rare, much less frequent than those with more than 4 inches. It is to be expected, therefore, that in general precipitation in Ohio is quite sufficient for wheat, and there is more likely to be an oversupply than an undersupply of moisture. On the other hand, the snowfall is not heavy, much of the winter precipitation being rain, the ground is not long nor deeply covered, subzero (F.) temperatures are frequent, temperatures are subject to large fluctuations, and rise to high values early in the summer. Hence it is to be expected, perhaps, that the condition and yield will show more relation to temperature than to precipitation, Other factors which may have more or less influence on yield are sunshine, snowfall, snow cover, the extremes of temperature, the frequency and duration of very low or very high temperatures, the distribution of precipitation, the length and severity of droughts, etc. There are innumerable combinations of these factors which may affect the growth of the wheat plant, but the difficulty of securing the data for all these factors over considerable areas and periods, and of expressing them in simple numerical form, are great, if not insurmountable, and for the most part the following study has necessarily been confined to temperatures and precipitation, either as monthly or 10-day means and totals, with a few other factors introduced in some of the equations.

There is another set of contributing influences, not wholly disconnected with the weather, but which require separate treatment, and for which extensive data are not available. This includes injury by hessian fly or other insects, by rust, by severe storms, as of hail, and the loss by storms after the crop is practically made or actually cut. As the injury by hessian fly is largely a function of the date of planting and the weather, and that by rust a function of the condition of the seed and the weather, all factors are to a certain extent included in the weather factors.

[3] For this particular form of development I am indebted to a memorandum by Prof. C. F. Marvin, Chief of Bureau.
[4] Moore, H. L. Forecasting the Yield and the Price of Cotton.

AVERAGE CONDITIONS FOR THE STATE.

Because of the length of the growing season and the multiplicity of factors that may be expected to affect the growth of winter wheat, and because of the necessity of keeping the number of terms in the equations small, an attempt was first made to consider the progress of the crop from month to month, or for a few months at a time, rather than for the whole season from planting to harvesting, by utilizing the "condition" reports of the Bureau of Crop Estimates. These reports give the estimated condition of the crop on the first day of December, April, May, and June of each season, and the report of April 1 gives the percentage of the crop abandoned. The "condition," as technically used by the Bureau of Crop Estimates, is expressed as a percentage of the "normal." The "normal" is represented by 100, and is defined as "a condition of perfect healthfulness, unimpaired by drought, hail, insects, or other injurious agencies," and is something better than an average but not a perfect or best possible condition. These condition reports are available for the 26-year series, 1893 to 1918.

The first equation was an attempt to determine the influence of the fall weather upon the growth of the wheat plant by comparing the condition, as reported on December 1, with the temperature, precipitation, and sunshine data of September, October, and November. As an illustration of the method used in this and subsequent calculations, the data, the derived values, and the final equation are given in Table 1 below.

TABLE 1.—*Data for calculation of equation expressing condition of winter wheat in Ohio on Dec. 1.*

1	2	3	4	5	6	7	8	9	10	11	26	27
Year.	Reported condition Dec. 1 (per cent).	Departure.	Temperature, October-November (°F.).	Departure.	Precipitation, September (inches).	Departure.	Precipitation, October-November (inches).	Departure.	Sunshine, October-November (per cent).	Departure.	Calculated condition (per cent).	Y—Y¹ (per cent).
	Y	y	x_1	D_1	x_2	D_2	x_3	D_3	x_4	D_4	Y¹	
1893	92	+ 4	46	− 1	1.6	−1.1	6.3	+1.3	44	− 4	89	− 3
1894	93	+ 5	46	− 1	3.3	+0.6	4.2	−0.8	54	+ 6	85	− 8
1895	74	−14	44	− 3	1.7	−1.0	5.3	+0.3	52	+ 4	81	+ 7
1896	101	+13	47	0	5.1	+2.4	3.8	−1.2	46	− 2	91	−10
1897	......	......	50	+ 3	0.8	−1.9	7.2	+2.2	51	+ 3	......	......
1898	102	+14	46	− 1	2.6	−0.1	6.9	+1.9	40	− 8	93	− 9
1899	83	− 5	50	+ 3	2.7	0	3.8	−1.2	47	− 1	89	+ 6
1900	86	− 2	51	+ 4	1.8	−0.9	6.1	+1.1	48	0	92	+ 6
1901	75	−13	46	− 1	2.9	+0.2	2.2	−2.8	52	+ 4	79	+ 4
1902	98	+10	52	+ 5	4.6	+1.9	4.9	−0.1	42	− 6	99	+ 1
1903	80	− 8	46	− 1	1.5	−1.2	4.7	−0.3	51	+ 3	82	+ 2
1904	76	−12	46	− 1	2.0	−0.7	1.9	−3.1	46	− 2	79	+ 3
1905	98	+10	46	− 1	2.9	+0.2	6.2	+1.2	46	− 2	90	− 8
1906	97	+ 9	47	0	2.9	+0.2	5.8	+0.8	39	− 9	93	− 4
1907	84	− 4	44	− 3	3.9	+1.2	4.7	−0.3	52	+ 4	86	+ 2
1908	62	−26	48	+ 1	0.6	−2.1	2.3	−2.7	60	+12	75	+13
1909	95	+ 7	49	+ 2	1.8	−0.9	4.8	−0.2	54	+ 6	86	− 9
1910	91	+ 3	46	− 1	4.0	+1.3	6.1	+1.1	44	− 4	93	+ 2
1911	83	− 5	46	− 1	4.9	+2.2	7.9	+2.9	38	−10	101	+18
1912	95	+ 7	48	+ 1	3.1	+0.4	3.5	−1.5	58	+10	84	−11
1913	99	+11	49	+ 2	2.4	−0.3	6.9	+1.9	41	− 7	95	− 4
1914	94	+ 6	50	+ 3	1.4	−1.3	4.9	−0.1	51	+ 3	87	− 7
1915	85	− 3	50	+ 3	4.5	+1.8	5.1	+0.1	56	+ 8	93	− 8
1916	87	− 1	48	+ 1	2.6	−0.1	4.2	−0.8	61	+13	83	− 4
1917	83	− 5	43	− 4	1.9	−0.8	5.3	+0.3	41	− 7	84	+ 1
Sum	2,113	+ 1	1,184	+ 9	67.5	0	125.0	0	1,214	+14	2,109	−20
Mean	88.0	......	47.4	......	2.70	......	5.00	......	48.5	......	87.9	−0.8

(Sunshine percentages are averages of data for Cincinnati, Cleveland, Columbus, and Toledo, Ohio, and Parkersburg, W. Va. Other data are Ohio State means.)

Columns 12 to 25 contain squares and products of departures, from which the following sums are obtained:

$\Sigma D_1^2 = +131$
$\Sigma D_2^2 = +37.44$
$\Sigma D_3^2 = +58.24$
$\Sigma D_4 = 1,048$

$\Sigma D_1 D_2 = -1.3$
$\Sigma D_1 D_3 = +2.7$
$\Sigma D_1 D_4 = +63$

$\Sigma D_2 D_3 = +4.01$
$\Sigma D_2 D_4 = -52.8$
$\Sigma D_3 D_4 = -137.2$

$\Sigma D_1 y = +128$
$\Sigma D_2 y = +109.2$
$\Sigma D_3 y = +173.6$
$\Sigma D_4 y = -627$

Normal equations:

$$\begin{aligned} 128 &= 131b_1 - 1.3\,b_2 + 2.7\,b_3 + 63\,b_4 \\ 109.2 &= -1.3b_1 + 37.44b_2 + 4.01b_3 - 52.8b_4 \\ 173.6 &= 2.7b_1 + 4.01b_2 + 58.24b_3 - 137.2b_4 \\ -627 &= 63\,b_1 - 52.8\,b_2 - 137.2\,b_3 + 1{,}048\,b_4 \end{aligned}$$

Solving: $b_1 = 1.1$, $b_2 = 2.3$, $b_3 = 2.1$, $b_4 = -0.3$.
Substituting in equation (3): $a = 33.8$.
Substituting in equation (1):

$$Y = 33.8 + 1.1x_1 + 2.3x_2 + 2.1x_3 - 0.3x_4$$

Substituting the values of x_1, x_2, x_3, x_4 for individual years in this equation, we get the calculated condition of December 1 as given in column 26. A comparison of the calculated and reported values shows differences ranging from 1 to 18 per cent, with an average difference, disregarding sign, of 6.3 per cent. The square root of the mean of the squares of the differences, which I have called the scatter, is 7.5, while the actual variability of Y, as expressed by a similar quantity, the standard deviation, is 9.8, showing a considerable improvement over chance results. We seem to have here as nearly as they can be expressed by monthly averages for a whole State the main factors on which the December 1 condition depends. Comparing the values of $b_1\sigma_1$, $b_2\sigma_2$, $b_3\sigma_3$, $b_4\sigma_4$, we would say that the precipitation of October and November is the most important of the four factors, but that all have considerable influence, and would conclude that a wet September, October, and November, the two latter being also warm and cloudy, are the factors favoring a high condition.

Making similar calculations for the change of condition from December 1 to April 1, for the percentage abandoned on April 1, and for the change of condition from May 1 to June 1, we get the results shown in Table 2. It was necessary to omit the change from April 1 to May 1, because the condition figures are vitiated for comparative purposes by the omission of the abandoned areas after April 1. In this table the values in each row are relative to each other, but values from different rows have no relative significance. It will be noted also that values in the third row, expressing the percentage abandoned, should be opposite in sign to those expressing condition.

TABLE 2.—*Relative values of weather factors affecting the condition of winter wheat in Ohio on various dates.*

[Values of $b\sigma$.]

	Condition on Dec. 1.	Change of condition, Dec. 1 to Apr. 1.	Percentage abandoned Apr. 1.	Change of condition, May 1 to June 1.
Temperature (° F.):				
October to November	2.52	....	....	....
October to February	....	....	5.71	....
December to February	....	1.80	....	....
March	....	4.31	−8.21	....
May	....	....	....	−1.5
Precipitation (inches):				
September	2.76	....	....	....
September to November	....	....	−0.34	....
October to November	3.15	....	....	....
December to February	....	−3.99	7.22	....
March	....	−0.65	....	....
May	....	....	....	5.4
Sunshine:				
October to November	−1.95	....	−2.48	....
March	....	0.81	....	....
May	....	....	....	1.4
Snowfall, March (inch)	....	−0.89	....	....
σy	9.8	9.6	12.3	7.6
S	7.5	8.1	9.2	5.8

The actual relation between the computed and recorded yields, as shown by the differences between σy and S, is not very close, but is perhaps sufficient to justify the use of these equations as preliminary indications of the important factors to be considered. We should say that they indicate, as factors favoring a high condition, (*a*) temperatures above normal from October to March, inclusive, especially in March, and temperatures below normal in May; (*b*) precipitation above normal from September to November, inclusive, below normal from December to March, inclusive, and above again in May. They indicate that the variations in the amount of sunshine in October, November, March, and May, and the amount of snowfall in March, are relatively unimportant factors.

Selecting what appear to be the important factors in the above table, and combining them in one equation for the entire season, and correlating them with the yield instead of the condition, we get the following values for the quantity $b\sigma$:

Temperature (°F.):	
March	1.22
October to February	0.67
Precipitation (inch):	
December to February	−0.42
September to November	−0.35
May	−0.05

It will be noticed that the precipitation of September to November and of May is negative instead of positive as in the preceding table where estimated conditions were being considered. The negative value for May is confirmed and increased in the subsequent calculations. The inconsistency may perhaps be explained on the assumption that a heavy rainfall in May will produce an amount of growth which gives the plant an apparently high condition on June 1, but is not actually conducive to a high yield of seed. In other plants it has been found that a vigorous growth of foliage prevents a heavy yield of seed. The data for this last computation are for the 64-year period, 1855 to 1918, and the results should be more dependable than those for the shorter period, but the value of the derived equation ($Y=-3.0+0.27x_1+0.26x_2-0.15x_3-0.19x_4-0.04x_5$) as a forecasting equation is disappointingly small, since S is found to equal 3.42, and σy equals 3.70.

The yield data for this long period show a progressive increase independent of the weather, and to eliminate this secular variation, the next calculation was based on the departures from successive 10-year means, of which the value under consideration was the sixth. The method of partial correlation was used, and factors for April and June introduced, in addition to those above, with the results shown in Table 3.

TABLE 3.—*Partial correlation coefficients. Yield of winter wheat in Ohio correlated with monthly temperature and precipitation data.*

[63 years, 1855 to 1918 (omitting 1900).]

	Temperature (° F.).						Precipitation (inch).	
	October to November.	December to February.	March.	April.	May.	June.	May.	June.
1. (6th order)	+0.07	+0.10	+0.29	....	−0.26	+0.19	−0.16	+0.07
2. (4th order)	....	....	+0.30	−0.01	−0.28	+0.18	−0.10	....

After the process of elimination accomplished by the preceding equations, we may consider that we have here combined all the most important weather factors affecting the yield of winter wheat in Ohio in so far as they can be expressed in values for calendar months. As partial correlation coefficients these are all too small to be of much significance, and show that, at least, some of the most important factors controlling yield have not been included, or are not properly expressed in monthly means and totals. This was confirmed by the calculation of a final equation using the four most important of the above-named factors, which gives a scatter of 3.34, as compared with the standard deviation of 3.70; some slight improvement over the former equation, but of little value. We have then reached the negative conclusion that for the State of Ohio as a whole there are no monthly weather values vitally affecting the yield of winter wheat, but that on the whole a warm March and June and a cool and dry May are favorable. All other temperature and precipitation values may be wholly disregarded.

The failure of the yield to show definite response to monthly values is thought to be due to two principal causes: First, the diversity of conditions and differing stages of growth reached in different parts of the State, giving rise to opposite effects. Second, shorter periods at critical stages of growth may have more important effects than periods of a month, and two such periods occurring in the same month may have opposite effects. Accordingly the next step was to confine the investigation to smaller areas and to 10-day periods.

FULTON COUNTY.

For this purpose, Fulton County, in northwestern Ohio, was first studied, the data being readily available in MONTHLY WEATHER REVIEW, SUPPLEMENT No. 2, for the 29-year period, 1883 to 1912. The results of the preliminary study using monthly data are shown in the accompanying Table 4.

TABLE 4.—*Partial correlation coefficients. Yield of winter wheat in Fulton County correlated with monthly temperature and precipitation data at Wauseon, Ohio.*

[30 years, 1883-1912.]

	Temperature (° F.).						Precipitation, December-February (inch).	Snowfall, March (inch).
	October-November.	December-February.	March.	April.	May.	June 1-15.		
1. (4th order)	−0.11	0.30	0.40	....	−0.30	....	−0.21	....
2. (5th order)	−0.11	0.37	0.37	−0.28	−0.15	....	−0.21	....
3. (6th order)	−0.10	0.37	0.38	−0.29	−0.17	0.11	−0.22	....
4. (6th order)	−0.06	0.42	0.04	−0.06	−0.26	....	−0.10	−0.61

The first three sets of coefficients agree very well with those for the State, showing that the December to February temperatures and the March temperatures should be above normal; the April and May temperatures below normal, and the winter precipitation below normal. But by introducing the snowfall of March, as shown in the fourth row, the influence of the March temperature practically disappears, as does that of the April temperature, and the March snowfall becomes much the most important factor, the only others of importance being the winter temperature and the May temperature. We will return to the consideration of this point after examining the influence of shorter periods during the active growing season.

The average date of ripening of winter wheat at Wauseon, Fulton County, is July 5, and the average date of blossoming is June 9. (MONTHLY WEATHER REVIEW, SUPPLEMENT No. 2). Counting back by 10-

day periods from July 5, nine periods take to April 6, about the beginning of active growth, and one of these periods, June 5 to 14, will have as its fifth day the average date of blossoming, June 9. Using the mean temperatures and the total precipitation for these periods, and omitting other factors for the present, we may briefly summarize the results of numerous calculations in lines 1 to 8, inclusive, of Table 5, expressing the relationships by means of partial correlation coefficients.

TABLE 5.—*Partial correlation coefficients (sixth order). Yield of winter wheat in Fulton County, Ohio, correlated with temperature and precipitation data at Wauseon, Ohio.*

[30 years, 1883–1912.]

	Temperature (° F.).								
	Apr. 6–15. t_1	Apr. 16–25. t_2	Apr. 26–May 5. t_3	May 6–15. t_4	May 16–25. t_5	May 26–June 4. t_6	June 5–14. t_7	June 15–24. t_8	June 25–July 4. t_9
	Tillering.		Jointing.			In boot.	Heading, filling, ripening, blossoming, in milk.		
1	....	....	−0.59	0.00	....	−0.24	+0.23	....	....
2	....	....	....	....	....	....	....	+0.49	....
3	....	....	....	....	+0.20	....	....	....	....
4	....	....	....	....	....	....	....	....	+0.01
5	....	....	....	....	....	....	....	....	....
6	....	+0.42	....	....	....	....	....	....	....
7	−0.05	....	....	....	....	....	....	....	....
8	....	+0.65	−0.89	....	....	−0.21	....	+0.62	....
9	....	....	....	....	....	....	....	....	....
10	....	+0.58	−0.62	....	....	....	....	+0.67	....
11	....	....	....	....	....	....	....	....	....

	Precipitation (inch).								Temperature March (° F.). tm	Snowfall, March (inch). S
	Apr. 16–25. p_2	Apr. 26–May 5. p_3	May 6–15. p_4	May 16–25. p_5	May 26–June 4. p_6	June 5–14. p_7	June 15–24. p_8	June 25–July 4. p_9		
	Tillering.	Jointing.			In boot.	Heading, filling, ripening, blossoming, in milk.				
1	....	....	−0.09	....	−0.33	+0.25	....	....	....	....
2	....	−0.13	....	....	....	....	....	....	....	....
3	....	....	....	+0.09	....	....	....	....	....	....
4	....	....	....	....	....	....	....	+0.30	....	....
5	....	....	....	....	....	....	−0.11	....	....	....
6	....	....	....	....	....	....	....	....	....	....
7	....	....	....	....	....	....	....	....	....	....
8	....	....	....	....	−0.24	+0.37	....	+0.36	....	....
9	....	....	....	....	....	....	....	....	+0.30	....
10	....	....	....	....	....	+0.30	....	+0.35	+0.03	−0.56
11	−0.26	....	....	....	....	....	....	....	....	....

After completing the calculation of the seven coefficients given in line 1, the method, it will be observed, was to drop the least important of these, and substitute one or two others, and determine their coefficients of the sixth order, without completing the calculation for the remaining terms, until the final computation in line 8. In this way all the quantities were introduced into a correlation calculation of seven terms, most of which had previously indicated their importance. Hence it is thought that fortuitous relations and those apparent relations resulting from large correlations among the weather factors themselves have been largely eliminated.

Several features of these coefficients should be noted. In the first place, three of those in line 8 are much larger than any in Table 3 or than any in Table 4, except the March snowfall. They are of such magnitude as to leave little doubt of a definite connection with the yield. In the second place, the sudden reversal of sign in adjacent 10-day periods is to be noted, especially that between t_2 and t_3, and between p_6 and p_7. Do these correspond to reality? To test this matter a regression equation was formed including only the three factors, t_2, t_3, and t_8. The resulting equation was

$$Y = 0.53 + 0.14x_2 - 0.51x_3 + 0.52x_8 \qquad (6)$$

in which the yield is made to depend absolutely upon these three temperature values. The resulting scatter is 3.54, while the standard deviation is 4.59, which is 1.3 times the scatter.

A comparison of the reported and calculated yields by individual years is shown in curve A, figure 1. In 16 of the 30 years the difference is 2 bushels or less. The only large differences are in the years 1899, 1900, and 1912, when there were extremely small yields, which in 1899 and 1900 were due to extensive injury by hessian fly. The average difference, disregarding sign, is 2.7, and omitting the three bad years, the average is 2.1. The mean of the reported yields is 15.6 and of the calculated yields is 16.2. We can not doubt that these three 10-day temperature periods have had a marked influence on the yield in Fulton County, and that the first and third periods should be warm, and the second cool. The coefficients for p_6 and p_7, being much smaller, a similar study had not been made for them, but the opposite signs persist in all combinations.

What may be the physical explanation of these facts? Mr. F. A. Welton, of the Ohio Agricultural Experiment Station, has kindly given me the following approximate dates of occurrence of various stages in the growth of winter wheat at Wooster, in northeastern Ohio.

Tillering, April 1 to May 1.
Jointing, April 25 to May 20.
In boot, May 20 to May 28.
Heading, May 28 to June 8.
Blossoming, one or two days after heading.
Filling, June 10 to June 20.
In milk, June 20 to June 30.
Ripening, June 30 to July 12.

Now the record at Wauseon, Fulton County, gives July 5 as the average date of ripening, which corresponds very closely with that given for Wooster, and June 9 as the date of blossoming, which is not more than five days later than at Wooster. Assuming that the earlier dates also show approximately the same difference, we conclude that the period represented by t_2 and t_3, when the sign of the temperature influence changes, is the time when tillering has been completed and jointing begins. For the completion of the spreading process relatively warm April weather is needed, but at the beginning of the rapid growth in height of the plant cool weather is beneficial. The large positive value of t_8 indicates the importance of warm weather during the period of filling. Similarly, the values for p_6 and p_7 indicate that the 10 days preceding heading and blossoming, when the wheat is said to be in the boot, should be dry, but that during the heading and blossoming period rain is beneficial. Finally, the coefficient for p_9 indicates that the yield is increased when there is more than the average rainfall during the 10 days before ripening, when the grain is in the milk stage. In all cases, however, the rainfall coefficients are less important and less conclusive than the temperature coefficients.

Returning now to the consideration of the influence of March temperature and snowfall, we substitute the March temperature in the place of t_8, of line 8, Table 5, and we find a coefficient of +0.30. But there is, as might be expected, a large negative relation between the March temperature and snowfall, expressed by the total correlation coefficient, −0.60. If now we retain the March temperature in place of t_6 and introduce the March snowfall, instead of p_6 we get the results shown in line 10, Table 5; that is to say, the relation of yield to March temperature practically disappears, but a large negative relation to snowfall appears of nearly equal importance with the temperature relations of April 16 to May 5 and June 15 to 24. Introducing the March snowfall, then, as a fourth factor, we get, instead of equation 6,

$$Y = -9.4 + 0.25t_2 - 0.31t_3 + 0.47t_8 - 0.48s \qquad (7)$$

The resulting scatter is 3.04, as compared with the standard deviation of 4.59, a ratio of 1 to 1.5, showing an improvement in accuracy over equation 6. Curve B of figure 1 gives the comparison by years. The average deviation is 2.3, and 20 of the years show differences of 2 or less. The largest deviation is 7 in the hessian-fly year of 1900, but in the 2 other years of very light yield, 1899 and 1912, for which a large discrepancy was shown in curve A, the addition of the snowfall term results in a close agreement. The average of the reported yields and the calculated yields is exactly the same, 15.6 bushels, indicating that the equation might be used to determine with accuracy the average yield over a series of years, as in a case where weather data were available but yield data lacking.

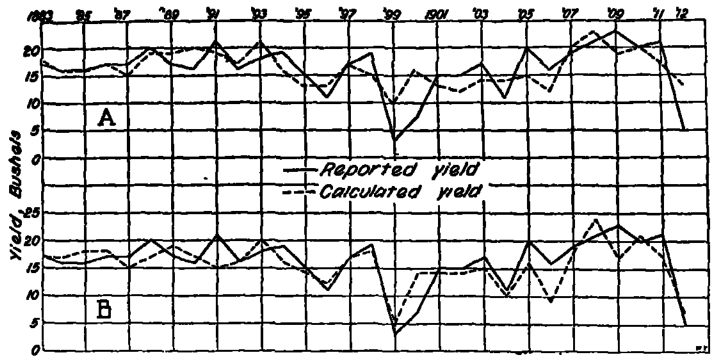

FIG. 1.—Comparison of reported and calculated yield of wheat in Ohio, showing calculated values (B) with and (A) without the snowfall term.

Considering the fact that we are using averages, and are therefore necessarily disregarding the fact that the time of occurrence of those stages that are particularly susceptible to weather influences will differ in different years, and in the same year in different fields, and considering that we are also disregarding unusual and extraneous factors, which are occasionally of large importance, it would seem that we get surprisingly close conformity. A careful application of such an equation to an individual plot, adjusting the temperature periods to the actual growth stages each year, should give interesting and valuable results.

CENTRAL OHIO.

For further and independent study, three counties in central Ohio were selected, namely, Madison, Franklin and Pickaway. The temperature data used are for the Weather Bureau station at Columbus, Franklin County; the precipitation data are the averages for all the stations in the three counties reporting rainfall, the number varying in individual years from 8 to 18; the yield data are the averages for the three counties. To check back with the State data an equation was first evaluated, using the important temperature and precipitation data of autumn, winter, and spring. This gave the following values for $b\sigma$:

Temperature (°F.):	
October–November	0.33
December–February	0.78
March	0.94
Precipitation, December–February (inch)	−0.45

The results agree with the State results, showing the temperature of March as the most important factor, and the temperature of the winter next in importance, but the scatter is 4.43 and the standard deviation is 4.56, showing that these factors are not important in determining yield.

Next, numerous partial correlations were worked out, using the most important of these factors together with the ten-day temperature and precipitation data during the season of active growth. The resulting values are shown in Table 6.

TABLE 6.—*Partial correlation coefficients. Sixth order. Yield of winter wheat in central Ohio correlated with temperature and precipitation.*

[Data for 28 years, 1889–1917 (omitting 1900).]

	Temperature (°F.).								
	Apr. 1-10 t_1	Apr. 11-20 t_2	Apr. 21-30 t_3	May 1-11 t_4	May 12-21 t_5	May 22-31 t_6	June 1-10 t_7	June 11-20 t_8	June 21-30 t_9
	Tillering.		Jointing.		In boot.	Heading, filling, blossoming.		In milk.	Ripening.
1				−0.68	+0.02	−0.58	+0.43		
2				−0.58		−0.53	+0.45	0.00	
3				−0.69		−0.67	+0.51		
4									
5									
6									
7				−0.75		−0.64	+0.54		+0.34
8		−0.16	−0.10	−0.72		−0.70	+0.54		+0.47
9	−0.15								

	Precipitation (inches).						Snowfall, March (inches).	Temperature (°F.).		
	May 1-11 p_4	May 12-21 p_5	May 22-31 p_6	June 1-10 p_7	June 11-20 p_8	June 21-30 p_9		December to February.	March.	April.
	Jointing.	In boot.	Heading, filling, blossoming.		In milk.	Ripening.				
1							+0.03	−0.19	+0.01	
2	+0.03							−0.17		−0.10
3		−0.43	−0.25				+0.02	−0.15		
4				−0.10						
5					−0.17					
6						−0.26				
7		−0.49	−0.11			−0.19				
8		−0.40								
9										

The insignificance of the December to February and the March temperatures as compared with 10-day temperature and rainfall later in the season is evident from the table, agreeing with the results for Fulton County. In contrast with the Fulton County result, the March snowfall is also unimportant. The average snowfall of this section for March is only 4 inches, with 15 out of 28 years having 2 inches or less, compared with an average fall in Fulton County of 7 inches, with only 5 out of 30 years having 2 inches or less. The unimportance of the snow in the central counties, therefore, is probably due to the small amount that falls and to the fact that it accordingly disappears quickly.

How closely do the important 10-day periods correspond in the two sections of the State? Welton[5] states that the season in central Ohio is from 5 to 7 days earlier than at Wooster, as given above, and we found about 5 days difference between Wooster and Fulton County, hence central Ohio should be from 10 to 12 days earlier than Fulton County. We notice, first, the absence of an important relation during the early jointing stage, but we do find indication of a negative relation in April corresponding, roughly, to t_3 of Table 5, and a very large negative relation during the first decade of May, which is also within the period of jointing. So we have substantial agreement here that during the period of rapid growth in height temperatures should be below normal. For the period of filling, t_8 in Table 5 and t_7 in Table 6, we have exact agreement, temperatures above normal are decidedly favorable. In central Ohio we find also that cool weather is highly important in the preceding period, that of blossoming, while Fulton County fails to show this relation definitely, but there is indication of a negative relation in the period just before blossoming. Warm weather is indicated for the period of ripening in central Ohio, while the corresponding relation is insignificant in the northern county. The two tables agree in showing a negative precipitation relation during the time when the grain is in the boot.

The regression equation embodying the five important factors in line 8, Table 6, is:

$$Y = 29.7 - 0.46x_1 - 0.58x_2 + 0.42x_3 + 0.33x_4 - 1.97x_5 \quad (8)$$

in which x_1 = temperature, May 1–11; x_2 = temperature, May 22–31; x_3 = temperature, June 1–10; x_4 = temperature, June 21–30; x_5 = precipitation, May 12–21.

A comparison of the yields computed from this equation with those reported by the Bureau of Crop Estimates is made in figure 2. The average deviation of the calculated from the reported yield is 2.3 bushels; there is only one difference as great as 5, and 18 of the 28 years

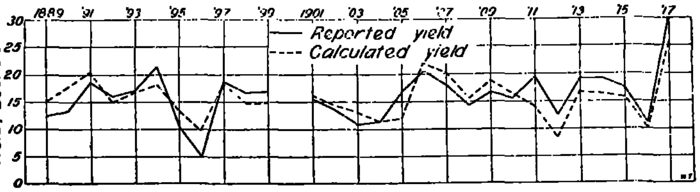

FIG. 2.—Comparison of reported and calculated yield of wheat in Ohio. (Computation from equation 8.)

have differences of 2 or less; the scatter is 2.75, while the standard deviation of the yield is 4.56, which is 1.7 times the scatter. In all but two of the years the change from the previous year is in the same direction for the reported and calculated values, showing the controlling influence of these weather elements, even when the resulting yield is not strictly proportional to them. The average calculated yield for the 28 years is 15.7 bushels, the same as the average of the yields actually reported.

Further to test the soundness of this equation it was applied unchanged to the data for 1918 and 1919, which were not included in the calculation of the equation. The data and calculated values are:

	Temperature (°F.).				Precipitation May 12–21 (inches).	Reported yield (bushels).	Calculated yield (bushels).
	May 1–11.	May 22–31.	June 1–10.	June 21–30.			
1918	62	75	72	69	1.81	18.7	7.1
1919	56	66	76	72	1.12	21.1	19.1

[5] In private letter, quoted above.

It will be noted that for 1919 there is a deviation of but 2 bushels, strongly confirming the value of the equation, while for 1918 the deviation is so great as to cast doubt upon its value. But the season of 1918 was not a normal one. By the 10th of June "wheat was from 10 to 14 days in advance of the normal season" (so stated in National Weather and Crop Bulletin, issued June 11, 1918; also indicated in the issues of June 4 and June 18), due to temperatures much above normal from May 1 to June 10, especially during the period, May 22–31, for which the average temperature, 75°, was 11° above normal. Hence the dates used do not correspond with the stages to which they have been applied, and the extremely high temperatures belong in the third period rather than the second. Advancing the dates, except for the period May 1–10, first, 10 days, and then the later ones 14 days, we have the following values:

Temperature (° F.):	
May 1–11	62
May 12–21	68
May 22–21	75
June 11–20	68
Precipitation, May 1–11 (inch)	0.92
Reputed yield (bushels)	18.7
Calculated yield (bushels)	14.9
Temperature (° F.):	
May 1–11	62
May 8–17	64
May 18–27	72
June 7–16	68
Precipitation, May 1–11 (inch)	0.92
Reported yield (bushels)	18.7
Calculated yield (bushels)	15.0

By this adjustment to the actual growth history the calculated yield is brought within a resonable deviation from the recorded yield, and this abnormal year offers further proof of a direct relation between the yield and the temperature and precipitation values at certain fixed stages in the growth of the plant.

CONCLUSIONS.

1. For the State of Ohio as a whole, a warm March and June and a cool and dry May are favorable conditions for a high yield of winter wheat. All other monthly temperature and precipitation values may be wholly disregarded.

2. There are certain 10-day periods during April, May, and June which appear to be the critical stages in the influence of the weather upon the yield of winter wheat in Ohio. The temperature and precipitation values during these 10-day periods largely determine yield.

3. In northern Ohio, represented by Fulton County, and in central Ohio, the weather should be cool during the jointing stage in the growth of the wheat plant, dry during the development of the boot, and warm while the head is filling.

4. In addition, it should be warm in Fulton County, during the last 10 days of the stooling process. It should be cool in central Ohio during blossoming and warm while the grain is ripening.

5. In Fulton County, and probably also in other counties that are subject to snows of sufficient depth to remain long on the ground, a heavy March snowfall is decidedly detrimental.

6. Because of the large influence of late May and June temperatures, earlier forecasts of yield can be of little value.

It is a pleasure to acknowledge my indebtedness to Prof. Charles F. Marvin, Chief of Bureau, and to Prof. J. Warren Smith for encouragement and suggestion in the prosecution of this study, and to the latter also for the loan of his valuable collection of data for Ohio, which I have used extensively and much of which I would not otherwise have obtained.